

KATEDRA OBECNÉ LINGVISTIKY
Posudek na magisterskou diplomovou práci

Autor DP: Bc. Vladimír Matlach

Studijní obor: obecná jazykověda a teorie komunikace

Název práce (česky/anglicky): Kvantitativně lingvistický software/Quantitative Linguistic Software

Vedoucí práce: Mgr. Radek Čech, Ph.D.

Oponent: Mgr. Martina Benešová, Ph.D.

Hodnoticí kritéria	Hodnocení
	A – F
Soulad názvu práce s obsahem, schopnost formulovat cíl a závěry práce	A
Teoretická podkladová koncepce, metodologie, terminologie	A
Schopnost analýzy excerpovaného materiálu/předmětu výzkumu a následné syntézy	
Interpretace a argumentace	B
Sekundární literatura, tvůrčí aplikace, schopnost kritického komentáře	
Struktura práce, výstavba textu, adekvátnost použitých stylistických prostředků, dodržení ortografické a gramatické normy	A
Bibliografický a poznámkový aparát, jednotný způsob citací	B
Výsledná známka	A

Hodnocení písemné verze práce:

Magisterská diplomová práce Bc. Vladimíra Matlacha si klade za cíl vytvoření softwaru QUITA, který by uživatelsky přátelskou formou umožnil zpracování přirozených i uměle vytvořených textů (software samotný umožňuje vytvoření náhodných textů, byť je „náhodnost“ založena prozatím na malém množství jednoduchých kritérií). Software je určen dle slov autora primárně pro kvantitativní lingvisty a nevyžaduje předchozí hlubší znalosti kvantitativní lingvistiky, matematiky či statistiky. Software v současnosti umožňuje také analýzu a zpracování řetězců DNA. Účelem je zjednodušení zpracovávání textů, úspora času a minimalizace chyb, ke kterým může dojít při „ručním zpracování“. Magisterská diplomová práce je autorem sestavena ve formátu uživatelské příručky.

V úvodu hodnocení si dovoluji rozporovat tvrzení autora, že pro použití software není třeba hlubší znalosti kvantitativní lingvistiky, neboť pro plnohodnotné a validní kvantitativně lingvistické i jiné experimenty, ve kterých je správně formulována hypotéza a cíle, je třeba znát nástroje minimálně kvantitativně lingvistické, lépe také alespoň základy matematiky i statistiky, pokud bude třeba je použít. Sám autor tento fakt aplikuje ve svých precizně strukturovaných příkladových studiích, viz např. kapitola 2.10.2 a aplikace standardní odchylky.

Autor ve své práci přehledně poukazuje na jednotlivé kroky zpracování textů, předkládá je a diskutuje jejich úskalí, viz např. úskalí tokenizace při odlišné definici jednotek. V této souvislosti je zajímavé umožnění zahrnutí či ignorace numerických či interpunkčních řetězců pro různé typy a cíle analýz.

Na úvod práce autor porovnává software QUITA se zdrojovým kódem vytvořeným pomocí volně dostupného statistického software R a zdůrazňuje úspornost a nepochybnost přátelskost prostředí QUITA. Je třeba podtrhnout, že software je vytvořen tak, aby měl uživatel možnost jeho úprav a doprogramování dle vlastních potřeb se základními znalostmi jazyka VB.NET, viz kapitola 4.

Mezi další nesporné výhody je třeba zařadit možnost exportu zpracovaných dat z různých stádií experimentu např. do Excelu a velká variabilita při přípravě grafického znázornění výstupů. Grafické modely je také možno exportovat.

Autor sám upozorňuje na slabé stránky současné verze software QUITA a předchází tak možným dotazům a nejasnostem.

V textu magisterské diplomové práce několikrát chybí odkazy na obrázky, např. v případě obrázků 12, 13, 14, 15.

Jak bylo již zmíněno výše, práce je velmi vhodně doplněna o přehledné a podrobné přehledové studie použití některých z indexů a jejich případného porovnání či vyhodnocení. V těchto miniprobλέmech navrhuji místo formy otázek používat standardní formát hypotéz problému.

Současná verze software QUITA umožňuje výpočet několika parametrů používaných v kvantitativní lingvistice, z rozhovorů s autorem je však patrné, že je plánováno rozšiřování možností QUITA. Považuji za vhodné zmínit, že software byl s kladným ohlasem prezentován na mezinárodní kvantitativně lingvistické konferenci QUALICO 2014.

Software QUITA i uživatelský manuál prezentovaný v této magisterské diplomové práci považuji za velmi zdařilý a po dalších úpravách a doplnění velmi dobře využitelný nejen v kvantitativně lingvistickém výzkumu. Práci doporučuji k obhajobě a po zodpovězení několika otázek a po následné diskusi navrhuji hodnocení A.

Náměty pro obhajobu (formou otázek):

1. Software QUITA je vytvořen v angličtině. Manuál k němu je prezentován ve formě magisterské diplomové práce v češtině. Je nebo bude manuál dostupný také v jazyce anglickém?
2. Jak bude probíhat údržba software a bude existovat nějaká uživatelská podpora v budoucnosti? Jaká rozšíření a podstatné úpravy a doplnění plánujete? Bude i nadále veřejně dostupný?
3. V manuálu (magisterské diplomové práci) postrádám seznam funkcí a parametrů, které současná verze QUITA umožňuje. Přednete prosím stručnou charakteristiku možností softwaru QUITA a ilustруйте konkrétní použití např. *entropie* v kvantitativní lingvistice.
4. V příkladu 2.10.2 je zmíněno menu Statistics. Z manuálu není patrné, kde je. Kde je a co umožňuje? Použití statistiky v kvantitativní lingvistice ilustруйте např. na u-testu, který zmiňujete v kapitolách 2.8.4 a 3.2.

Další poznámky k práci, kritické připomínky:

Datum a podpis oponenta práce:

Hodnocení průběhu obhajoby (prezentace, odpovědi na otázky a připomínky z posudků, diskuse v plénu) – návrh klasifikačního stupně:

Výsledné hodnocení – klasifikační stupeň:

Datum a podpis vedoucího komise pro obhajoby DP: